

Results for Abstract v.2

Sherly Boddu and Sam Hsu

4/21/2020

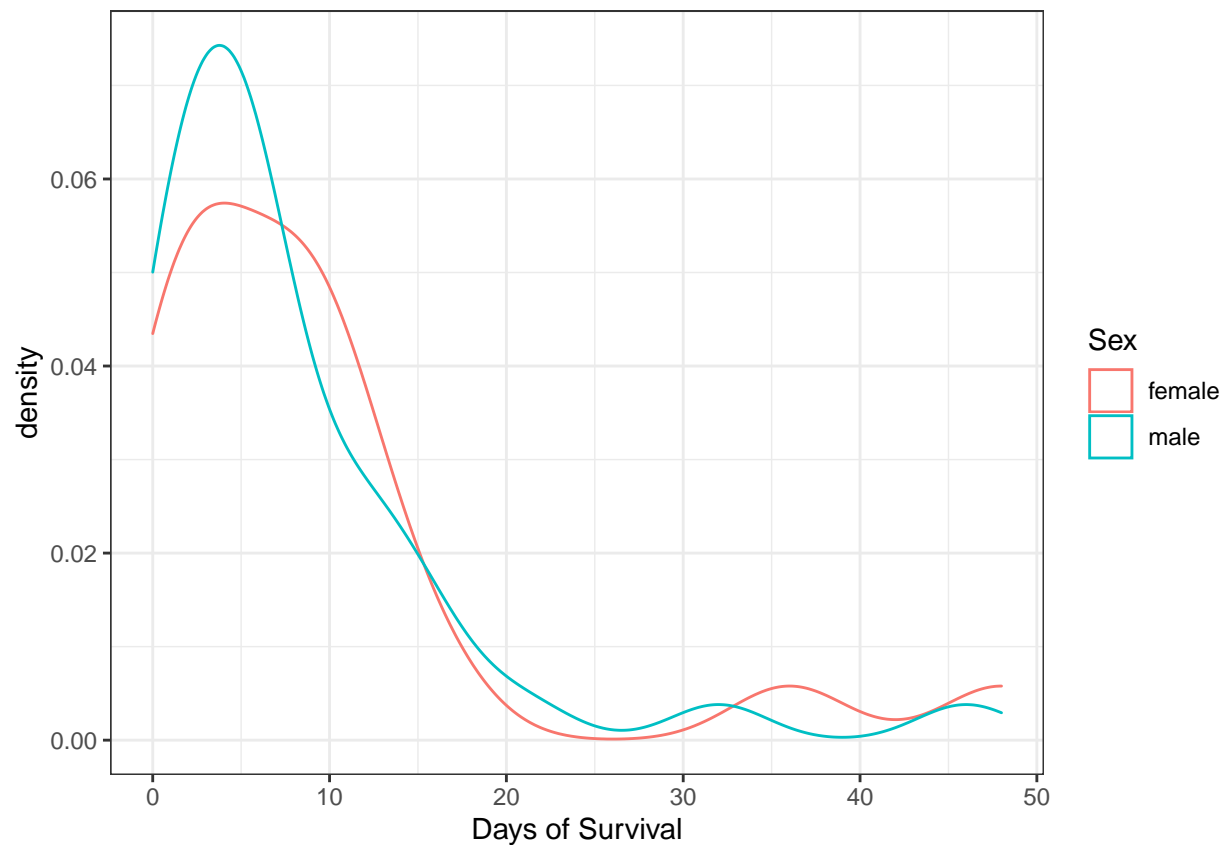
```
source("02_analysis/01_visualize.R")
source("02_analysis/02_k-m_stats.R")
```

Table 1: Descriptives	data_table01 (N = 3,324)
Sex:	
Male	1,428 (43.8%)
Female	1,830 (56.2%)
Age:	
Mean (SD)	45.03 (20.28)
Province:	
Gyeongsangbuk-do	1,221 (36.7%)
Gyeonggi-do	656 (19.7%)
Seoul	624 (18.8%)
Chungcheongnam-do	141 (4.2%)
Busan	129 (3.9%)
Gyeongsangnam-do	115 (3.5%)
Other	362 (10.9%)
Days to Resolution (Death or Release):	
Mean (SD)	21.17 (9.50)
Disease State:	
Deceased	65 (2.0%)
Isolated	1,622 (48.8%)
Released	1,637 (49.2%)

```
### density plots
```

```
data %>%
  filter(state == "deceased") %>%
  mutate(survival_days = as.numeric(survival_days, units = "days")) %>%
  ggplot(aes(x = survival_days, color = sex)) +
  geom_density(alpha = 0.3) +
  xlab("Days of Survival") +
  scale_color_discrete(name = "Sex") +
  theme_bw()
```

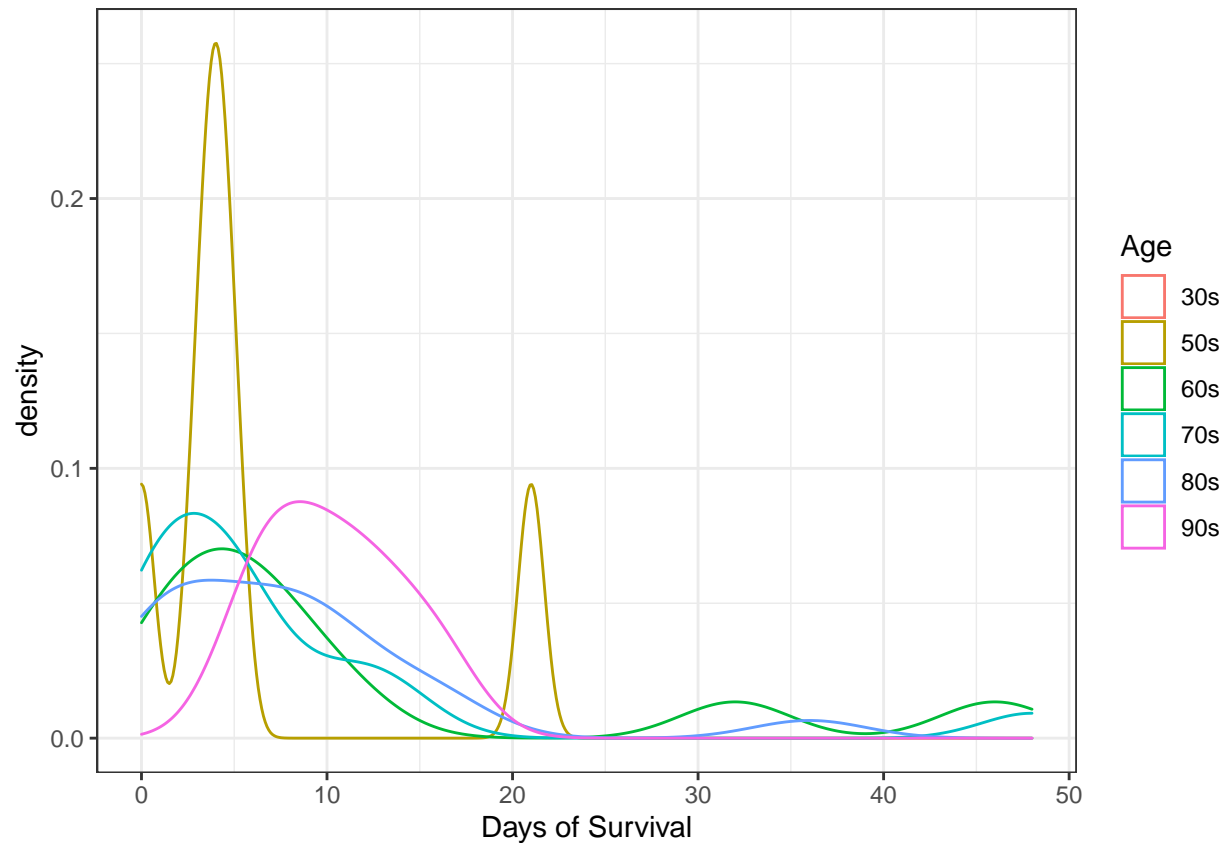
```
## Warning: Removed 6 rows containing non-finite values (stat_density).
```



```
data %>%
  filter(state == "deceased") %>%
  mutate(survival_days = as.numeric(survival_days, units = "days")) %>%
  ggplot(aes(x = survival_days, color = age_cat)) +
  geom_density(alpha = 0.3) +
  xlab("Days of Survival") +
  scale_color_discrete(name = "Age") +
  theme_bw()
```

```
## Warning: Removed 6 rows containing non-finite values (stat_density).
```

```
## Warning: Groups with fewer than two data points have been dropped.
```



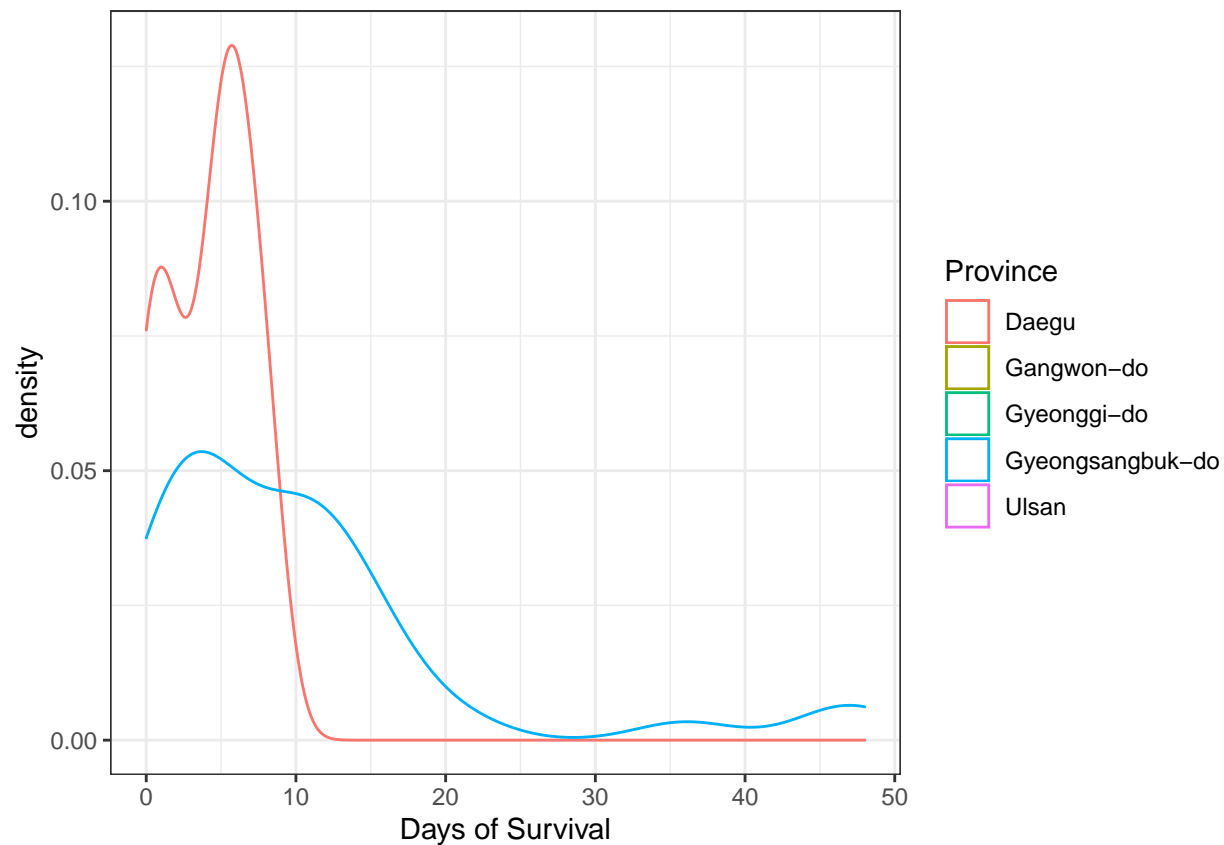
```
data %>%
  filter(state == "deceased") %>%
  mutate(survival_days = as.numeric(survival_days, units = "days")) %>%
  ggplot(aes(x = survival_days, color = province)) +
  geom_density(alpha = 0.3) +
  xlab("Days of Survival") +
  scale_color_discrete(name = "Province") +
  theme_bw()
```

```
## Warning: Removed 6 rows containing non-finite values (stat_density).
```

```
## Warning: Groups with fewer than two data points have been dropped.
```

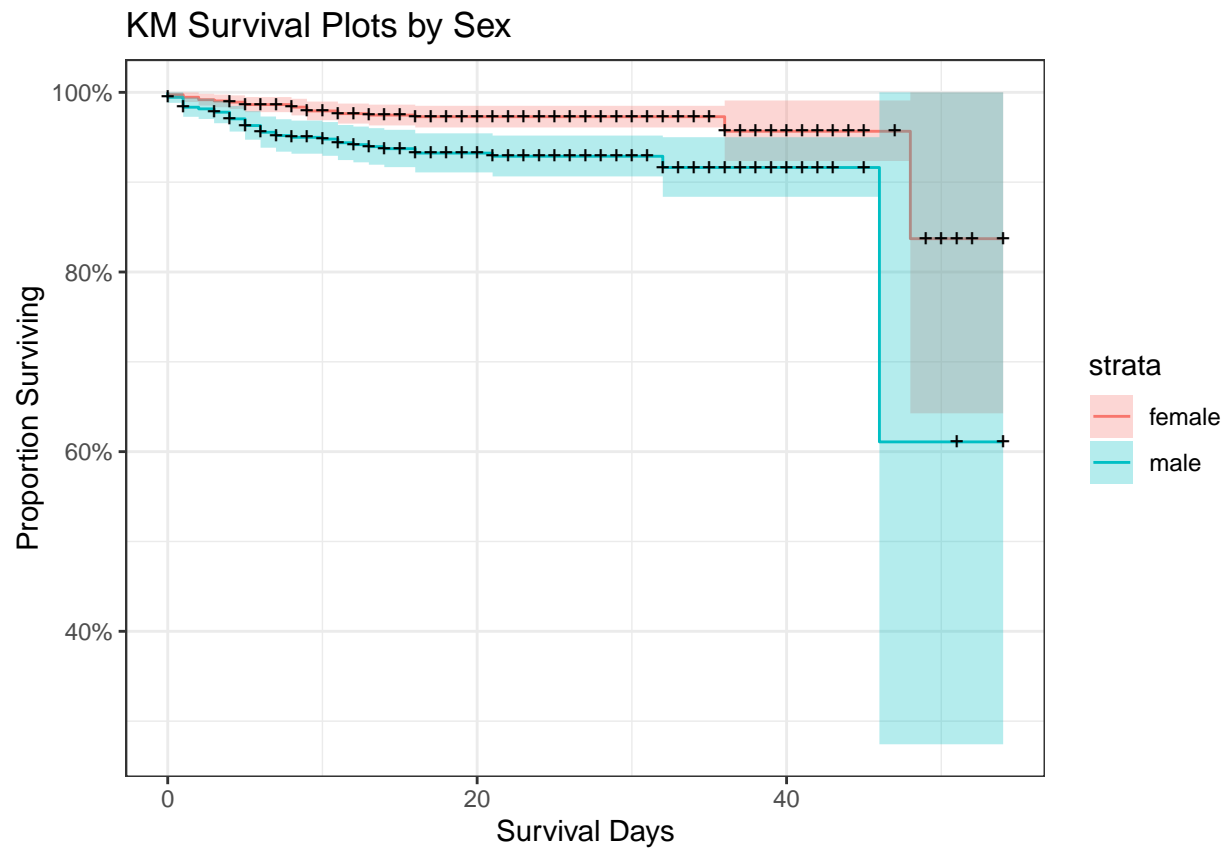
```
## Warning: Groups with fewer than two data points have been dropped.
```

```
## Warning: Groups with fewer than two data points have been dropped.
```

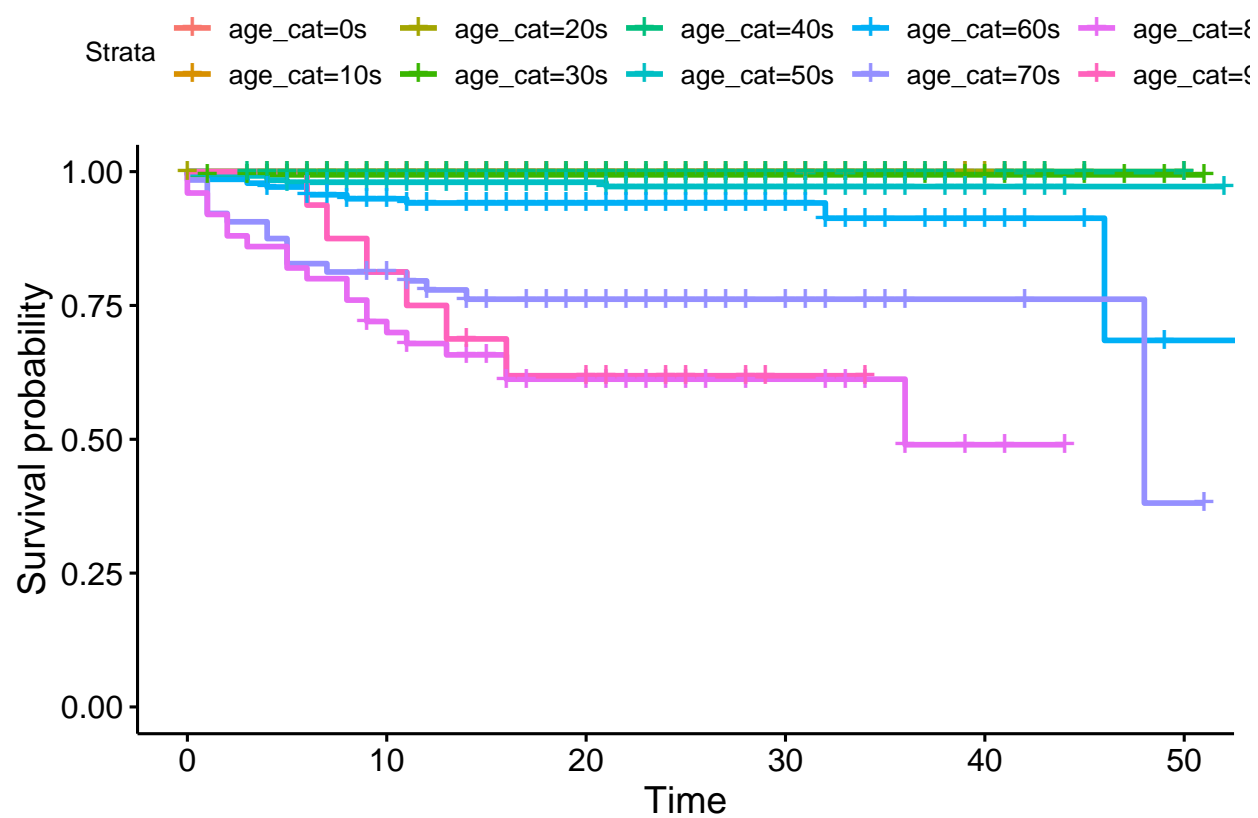


```
### K-M values and K-M curves
sex_surv <- survfit(Surv(survival_days, event, type = "right") ~ sex, data)
#age_surv_data <- data %>%
#  filter(age_cat %in% c("30s", "50s", "60s", "70s", "80s", "90s"))
age_surv <- survfit(Surv(survival_days, event, type = "right") ~ age_cat, data)
province_surv <- survfit(Surv(survival_days, event, type = "right") ~ province, data)

autoplot(sex_surv) +
  labs(x = "Survival Days", y = "Proportion Surviving", title = "KM Survival Plots by Sex") +
  theme_bw()
```

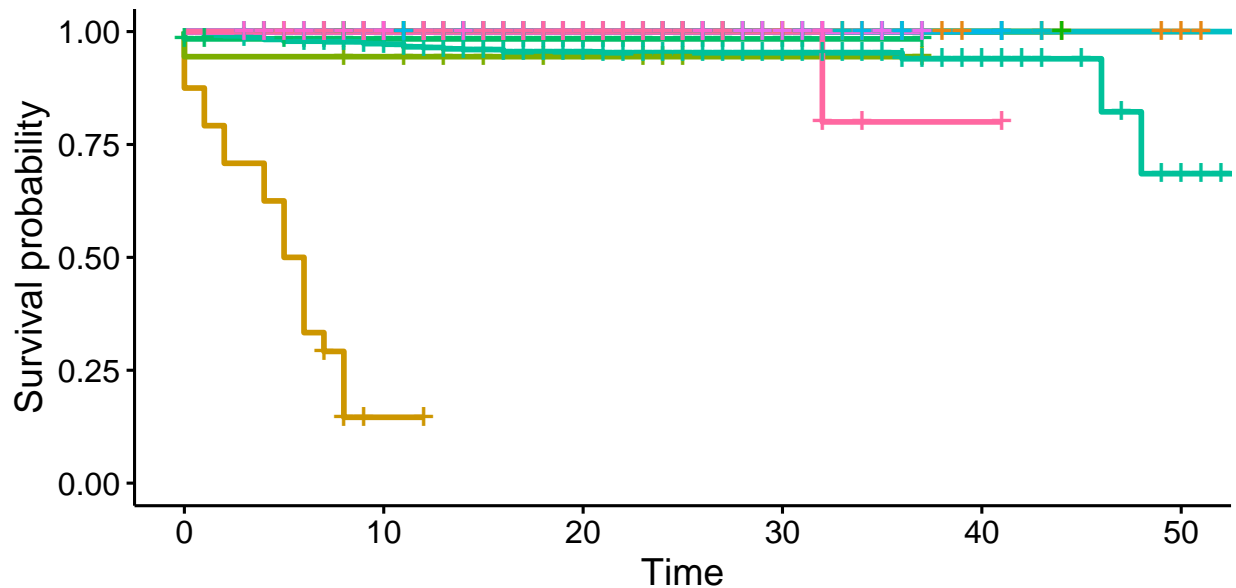


```
ggsurvplot(age_surv, data = data, conf.int = F)
```



```
ggsurvplot(province_surv, data = data, conf.int = F)
```

/vince=Chungcheongbuk-do + province=Gangwon-do + province=Gyeongsangnam-do +
 /vince=Chungcheongnam-do + province=Gwangju + province=Incheon +
 /vince=Daegu + province=Gyeonggi-do + province=Jeju-do +
 /vince=Daejeon + province=Gyeongsangbuk-do + province=Jeollabuk-do +



sex_surv

```
## Call: survfit(formula = Surv(survival_days, event, type = "right") ~
##       sex, data = data)
##
##      2048 observations deleted due to missingness
##              n events median 0.95LCL 0.95UCL
## sex=female 732      21    NA      NA      NA
## sex=male   544      38    NA      46      NA
```

age_surv

```
## Call: survfit(formula = Surv(survival_days, event, type = "right") ~
##       age_cat, data = data)
##
##      2050 observations deleted due to missingness
##              n events median 0.95LCL 0.95UCL
## age_cat=0s   17       0    NA      NA      NA
## age_cat=10s  52       0    NA      NA      NA
## age_cat=20s 324       0    NA      NA      NA
## age_cat=30s 167       1    NA      NA      NA
## age_cat=40s 194       0    NA      NA      NA
## age_cat=50s 251       6    NA      NA      NA
## age_cat=60s 139      10    NA      46      NA
## age_cat=70s  64      16   48      48      NA
## age_cat=80s  50      20   36      16      NA
## age_cat=90s  16       6    NA      13      NA
```

```
province_surv
```

```
## Call: survfit(formula = Surv(survival_days, event, type = "right") ~
##     province, data = data)
```

```
##
```

```
## 2048 observations deleted due to missingness
```

```
##           n events median 0.95LCL 0.95UCL
## province=Chungcheongbuk-do 22      0    NA      NA      NA
## province=Chungcheongnam-do 122     0    NA      NA      NA
## province=Daegu              24     20   5.5      4      8
## province=Daejeon            25      0    NA      NA      NA
## province=Gangwon-do         18      1    NA      NA      NA
## province=Gwangju            22      0    NA      NA      NA
## province=Gyeonggi-do        62      1    NA      NA      NA
## province=Gyeongsangbuk-do 751     36    NA     48     NA
## province=Gyeongsangnam-do  48      0    NA      NA      NA
## province=Incheon            56      0    NA      NA      NA
## province=Jeju-do             7      0    NA      NA      NA
## province=Jeollabuk-do        4      0    NA      NA      NA
## province=Jeollanam-do        2      0    NA      NA      NA
## province=Sejong             24      0    NA      NA      NA
## province=Seoul              54      0    NA      NA      NA
## province=Ulsan              35      1    NA      NA      NA
```

```
(sex_surv_lr <- survdiff(Surv(survival_days, event, type = "right") ~ sex, data))
```

```
## Call:
```

```
## survdiff(formula = Surv(survival_days, event, type = "right") ~
##     sex, data = data)
```

```
##
```

```
## n=1276, 2048 observations deleted due to missingness.
```

```
##
```

```
##           N Observed Expected (O-E)^2/E (O-E)^2/V
## sex=female 732      21     34.7      5.41     13.2
## sex=male  544      38     24.3      7.73     13.2
```

```
##
```

```
## Chisq= 13.2 on 1 degrees of freedom, p= 3e-04
```

```
(age_surv_lr <- survdiff(Surv(survival_days, event, type = "right") ~ age_cat, data))
```

```
## Call:
```

```
## survdiff(formula = Surv(survival_days, event, type = "right") ~
##     age_cat, data = data)
```

```
##
```

```
## n=1274, 2050 observations deleted due to missingness.
```

```
##
```

```
##           N Observed Expected (O-E)^2/E (O-E)^2/V
## age_cat=0s  17      0    0.804    0.804    0.818
## age_cat=10s 52      0    2.300    2.300    2.404
## age_cat=20s 324      0   14.892   14.892   20.071
## age_cat=30s 167      1    7.903    6.030    7.000
## age_cat=40s 194      0    8.962    8.962   10.609
## age_cat=50s 251      6   11.493    2.626    3.278
## age_cat=60s 139     10    7.015    1.270    1.466
## age_cat=70s  64     16    2.954   57.626   61.641
```



```

## age_cat=80s  50      20    1.995   162.469   168.902
## age_cat=90s  16       6    0.682    41.455    42.086
##
##  Chisq= 300  on 9 degrees of freedom, p= <2e-16
(province_surv_lr <- survdiff(Surv(survival_days, event, type = "right") ~ province, data))

## Call:
## survdiff(formula = Surv(survival_days, event, type = "right") ~
##   province, data = data)
##
## n=1276, 2048 observations deleted due to missingness.
##
##               N Observed Expected (O-E)^2/E (O-E)^2/V
## province=Chungcheongbuk-do  22         0   0.9192  9.19e-01  9.37e-01
## province=Chungcheongnam-do 122         0   6.2130  6.21e+00  7.04e+00
## province=Daegu              24        20   0.5299  7.15e+02  7.28e+02
## province=Daejeon            25         0   1.1116  1.11e+00  1.14e+00
## province=Gangwon-do         18         1   0.7134  1.15e-01  1.17e-01
## province=Gwangju            22         0   0.9791  9.79e-01  9.99e-01
## province=Gyeonggi-do        62         1   2.4968  8.97e-01  9.42e-01
## province=Gyeongsangbuk-do  751        36  35.8519  6.11e-04  1.57e-03
## province=Gyeongsangnam-do   48         0   2.4174  2.42e+00  2.53e+00
## province=Incheon            56         0   2.3731  2.37e+00  2.48e+00
## province=Jeju-do             7         0   0.3062  3.06e-01  3.09e-01
## province=Jeollabuk-do        4         0   0.1601  1.60e-01  1.61e-01
## province=Jeollanam-do        2         0   0.0747  7.47e-02  7.51e-02
## province=Sejong              24         0   1.1209  1.12e+00  1.15e+00
## province=Seoul               54         0   2.1447  2.14e+00  2.24e+00
## province=Ulsan               35         1   1.5879  2.18e-01  2.25e-01
##
##  Chisq= 740  on 15 degrees of freedom, p= <2e-16
(sex_surv_cox <- coxph(Surv(survival_days, event, type = "right") ~ sex, data))

## Call:
## coxph(formula = Surv(survival_days, event, type = "right") ~
##   sex, data = data)
##
##               coef exp(coef) se(coef)      z      p
## sexmale 0.9571      2.6041   0.2729  3.507 0.000454
##
## Likelihood ratio test=13.01 on 1 df, p=0.0003102
## n= 1276, number of events= 59
##   (2048 observations deleted due to missingness)
(age_surv_cox <- coxph(Surv(survival_days, event, type = "right") ~ approx_age, data))

## Call:
## coxph(formula = Surv(survival_days, event, type = "right") ~
##   approx_age, data = data)
##
##               coef exp(coef) se(coef)      z      p
## approx_age 0.089259  1.093364  0.008794 10.15 <2e-16
##
## Likelihood ratio test=140.5 on 1 df, p=< 2.2e-16

```

```
## n= 1067, number of events= 55
## (2257 observations deleted due to missingness)
(province_surv_cox <- coxph(Surv(survival_days, event, type = "right") ~ province, data))

## Warning in fitter(X, Y, istrat, offset, init, control, weights = weights, :
## Loglik converged before variable 1,2,3,4,5,6,7,8,9,10,11,12,13,14,15 ;
## coefficient may be infinite.

## Call:
## coxph(formula = Surv(survival_days, event, type = "right") ~
##       province, data = data)
##
##               coef exp(coef) se(coef)      z      p
## provinceChungcheongnam-do 4.916e-01 1.635e+00 1.660e+04 0.000 1.000
## provinceDaegu            2.320e+01 1.187e+10 1.596e+04 0.001 0.999
## provinceDaejeon          1.909e-01 1.210e+00 2.061e+04 0.000 1.000
## provinceGangwon-do        1.962e+01 3.327e+08 1.596e+04 0.001 0.999
## provinceGwangju           1.368e-01 1.147e+00 2.142e+04 0.000 1.000
## provinceGyeonggi-do       1.836e+01 9.412e+07 1.596e+04 0.001 0.999
## provinceGyeongsangbuk-do  1.922e+01 2.217e+08 1.596e+04 0.001 0.999
## provinceGyeongsangnam-do  5.074e-01 1.661e+00 1.755e+04 0.000 1.000
## provinceIncheon           8.330e-02 1.087e+00 1.858e+04 0.000 1.000
## provinceJeju-do           9.292e-02 1.097e+00 3.069e+04 0.000 1.000
## provinceJeollabuk-do      -3.115e-01 7.324e-01 4.813e+04 0.000 1.000
## provinceJeollanam-do      -6.265e-01 5.345e-01 8.102e+04 0.000 1.000
## provinceSejong            2.746e-01 1.316e+00 2.016e+04 0.000 1.000
## provinceSeoul              1.137e-03 1.001e+00 1.911e+04 0.000 1.000
## provinceUlsan             1.878e+01 1.429e+08 1.596e+04 0.001 0.999
##
## Likelihood ratio test=150.4 on 15 df, p=< 2.2e-16
## n= 1276, number of events= 59
## (2048 observations deleted due to missingness)
summary(sex_surv_cox)

## Call:
## coxph(formula = Surv(survival_days, event, type = "right") ~
##       sex, data = data)
##
## n= 1276, number of events= 59
## (2048 observations deleted due to missingness)
##
##               coef exp(coef) se(coef)      z Pr(>|z|)
## sexmale 0.9571      2.6041   0.2729 3.507 0.000454 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## sexmale      2.604      0.384      1.525      4.446
##
## Concordance= 0.617 (se = 0.032 )
## Likelihood ratio test= 13.01 on 1 df, p=3e-04
## Wald test = 12.3 on 1 df, p=5e-04
## Score (logrank) test = 13.25 on 1 df, p=3e-04
```

```
summary(age_surv_cox)
```

```
## Call:
## coxph(formula = Surv(survival_days, event, type = "right") ~
##       approx_age, data = data)
##
##      n= 1067, number of events= 55
##      (2257 observations deleted due to missingness)
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## approx_age 0.089259  1.093364 0.008794 10.15   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##              exp(coef) exp(-coef) lower .95 upper .95
## approx_age      1.093      0.9146      1.075      1.112
##
## Concordance= 0.906 (se = 0.015 )
## Likelihood ratio test= 140.5 on 1 df,  p=<2e-16
## Wald test               = 103 on 1 df,  p=<2e-16
## Score (logrank) test = 140.9 on 1 df,  p=<2e-16
```

```
summary(province_surv_cox)
```

```
## Call:
## coxph(formula = Surv(survival_days, event, type = "right") ~
##       province, data = data)
##
##      n= 1276, number of events= 59
##      (2048 observations deleted due to missingness)
##
##              coef exp(coef) se(coef)      z Pr(>|z|)
## provinceChungcheongnam-do 4.916e-01 1.635e+00 1.660e+04 0.000 1.000
## provinceDaegu             2.320e+01 1.187e+10 1.596e+04 0.001 0.999
## provinceDaejeon           1.909e-01 1.210e+00 2.061e+04 0.000 1.000
## provinceGangwon-do        1.962e+01 3.327e+08 1.596e+04 0.001 0.999
## provinceGwangju           1.368e-01 1.147e+00 2.142e+04 0.000 1.000
## provinceGyeonggi-do       1.836e+01 9.412e+07 1.596e+04 0.001 0.999
## provinceGyeongsangbuk-do  1.922e+01 2.217e+08 1.596e+04 0.001 0.999
## provinceGyeongsangnam-do  5.074e-01 1.661e+00 1.755e+04 0.000 1.000
## provinceIncheon           8.330e-02 1.087e+00 1.858e+04 0.000 1.000
## provinceJeju-do           9.292e-02 1.097e+00 3.069e+04 0.000 1.000
## provinceJeollabuk-do      -3.115e-01 7.324e-01 4.813e+04 0.000 1.000
## provinceJeollanam-do      -6.265e-01 5.345e-01 8.102e+04 0.000 1.000
## provinceSejong            2.746e-01 1.316e+00 2.016e+04 0.000 1.000
## provinceSeoul             1.137e-03 1.001e+00 1.911e+04 0.000 1.000
## provinceUlsan             1.878e+01 1.429e+08 1.596e+04 0.001 0.999
##
##              exp(coef) exp(-coef) lower .95 upper .95
## provinceChungcheongnam-do 1.635e+00 6.116e-01      0      Inf
## provinceDaegu             1.187e+10 8.426e-11      0      Inf
## provinceDaejeon           1.210e+00 8.262e-01      0      Inf
## provinceGangwon-do        3.327e+08 3.005e-09      0      Inf
## provinceGwangju           1.147e+00 8.722e-01      0      Inf
```

```

## provinceGyeonggi-do      9.412e+07  1.062e-08      0      Inf
## provinceGyeongsangbuk-do 2.217e+08  4.511e-09      0      Inf
## provinceGyeongsangnam-do 1.661e+00  6.021e-01      0      Inf
## provinceIncheon          1.087e+00  9.201e-01      0      Inf
## provinceJeju-do          1.097e+00  9.113e-01      0      Inf
## provinceJeollabuk-do     7.324e-01  1.365e+00      0      Inf
## provinceJeollanam-do     5.345e-01  1.871e+00      0      Inf
## provinceSejong           1.316e+00  7.599e-01      0      Inf
## provinceSeoul            1.001e+00  9.989e-01      0      Inf
## provinceUlsan            1.429e+08  6.997e-09      0      Inf
##
## Concordance= 0.796 (se = 0.023 )
## Likelihood ratio test= 150.4 on 15 df, p=<2e-16
## Wald test                = 162.6 on 15 df, p=<2e-16
## Score (logrank) test = 780.2 on 15 df, p=<2e-16
(surv_cox <- coxph(Surv(survival_days, event, type = "right") ~ sex + approx_age + province, data))

## Warning in fitter(X, Y, istrat, offset, init, control, weights = weights, :
## Loglik converged before variable 4,6,7,14 ; coefficient may be infinite.

## Call:
## coxph(formula = Surv(survival_days, event, type = "right") ~
##       sex + approx_age + province, data = data)
##
##               coef exp(coef) se(coef)      z      p
## sexmale          1.113e+00  3.042e+00  3.081e-01  3.611 0.000305
## approx_age       7.791e-02  1.081e+00  1.004e-02  7.759 8.54e-15
## provinceChungcheongnam-do 8.756e-01  2.400e+00  1.401e+04  0.000 0.999950
## provinceDaegu      2.152e+01  2.214e+09  1.281e+04  0.002 0.998660
## provinceGangwon-do -2.280e-01  7.961e-01  3.274e+04  0.000 0.999994
## provinceGyeonggi-do 1.824e+01  8.317e+07  1.281e+04  0.001 0.998864
## provinceGyeongsangbuk-do 1.903e+01  1.838e+08  1.281e+04  0.001 0.998815
## provinceGyeongsangnam-do 1.527e-01  1.165e+00  1.528e+04  0.000 0.999992
## provinceIncheon     3.474e-01  1.415e+00  1.660e+04  0.000 0.999983
## provinceJeju-do      1.026e+00  2.791e+00  3.956e+04  0.000 0.999979
## provinceJeollabuk-do 5.999e-01  1.822e+00  4.855e+04  0.000 0.999990
## provinceJeollanam-do 1.682e-01  1.183e+00  8.703e+04  0.000 0.999998
## provinceSeoul        4.755e-02  1.049e+00  1.571e+04  0.000 0.999998
## provinceUlsan        1.888e+01  1.582e+08  1.281e+04  0.001 0.998824
##
## Likelihood ratio test=238 on 14 df, p=< 2.2e-16
## n= 1067, number of events= 55
## (2257 observations deleted due to missingness)
summary(surv_cox)

## Call:
## coxph(formula = Surv(survival_days, event, type = "right") ~
##       sex + approx_age + province, data = data)
##
## n= 1067, number of events= 55
## (2257 observations deleted due to missingness)
##
##               coef exp(coef) se(coef)      z Pr(>|z|)
## sexmale          1.113e+00  3.042e+00  3.081e-01  3.611 0.000305 ***

```

```

## approx_age          7.791e-02  1.081e+00  1.004e-02  7.759  8.54e-15 ***
## provinceChungcheongnam-do  8.756e-01  2.400e+00  1.401e+04  0.000  0.999950
## provinceDaegu          2.152e+01  2.214e+09  1.281e+04  0.002  0.998660
## provinceGangwon-do     -2.280e-01  7.961e-01  3.274e+04  0.000  0.999994
## provinceGyeonggi-do    1.824e+01  8.317e+07  1.281e+04  0.001  0.998864
## provinceGyeongsangbuk-do 1.903e+01  1.838e+08  1.281e+04  0.001  0.998815
## provinceGyeongsangnam-do 1.527e-01  1.165e+00  1.528e+04  0.000  0.999992
## provinceIncheon        3.474e-01  1.415e+00  1.660e+04  0.000  0.999983
## provinceJeju-do        1.026e+00  2.791e+00  3.956e+04  0.000  0.999979
## provinceJeollabuk-do    5.999e-01  1.822e+00  4.855e+04  0.000  0.999990
## provinceJeollanam-do    1.682e-01  1.183e+00  8.703e+04  0.000  0.999998
## provinceSeoul          4.755e-02  1.049e+00  1.571e+04  0.000  0.999998
## provinceUlsan          1.888e+01  1.582e+08  1.281e+04  0.001  0.998824
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##               exp(coef) exp(-coef) lower .95 upper .95
## sexmale       3.042e+00  3.287e-01    1.663    5.565
## approx_age    1.081e+00  9.250e-01    1.060    1.103
## provinceChungcheongnam-do 2.400e+00  4.166e-01    0.000    Inf
## provinceDaegu  2.214e+09  4.517e-10    0.000    Inf
## provinceGangwon-do  7.961e-01  1.256e+00    0.000    Inf
## provinceGyeonggi-do  8.317e+07  1.202e-08    0.000    Inf
## provinceGyeongsangbuk-do 1.838e+08  5.440e-09    0.000    Inf
## provinceGyeongsangnam-do 1.165e+00  8.584e-01    0.000    Inf
## provinceIncheon  1.415e+00  7.065e-01    0.000    Inf
## provinceJeju-do   2.791e+00  3.583e-01    0.000    Inf
## provinceJeollabuk-do 1.822e+00  5.489e-01    0.000    Inf
## provinceJeollanam-do 1.183e+00  8.452e-01    0.000    Inf
## provinceSeoul     1.049e+00  9.536e-01    0.000    Inf
## provinceUlsan     1.582e+08  6.322e-09    0.000    Inf
##
## Concordance= 0.958 (se = 0.01 )
## Likelihood ratio test= 238 on 14 df,  p=<2e-16
## Wald test           = 89.48 on 14 df,  p=5e-13
## Score (logrank) test = 804.3 on 14 df,  p=<2e-16

```